

Rational engineering of multiple active sites in an ester hydrolase

Gerard Santiago^{†,○}, Mónica Martínez-Martínez^{‡,○}, Sandra Alonso[‡], Rafael Bargiela^{‡,¶}, Cristina Coscolín[‡], Peter N. Golyshin^{§,‡}, Víctor Guallar^{†,‡,*}, Manuel Ferrer^{‡,*}

[†]Barcelona Supercomputing Center (BSC), 08034 Barcelona, Spain

[‡]Institute of Catalysis, Consejo Superior de Investigaciones Científicas, 28049 Madrid, Spain

[¶]School of Chemistry, Bangor University, LL57 2UW Bangor, United Kingdom

[§]School of Biological Sciences, Bangor University, LL57 2UW Bangor, United Kingdom

[‡]Centre for Environmental Biotechnology, Bangor University, LL57 2UW, United Kingdom

[#]Institució Catalana de Recerca i Estudis Avançats (ICREA), 08010 Barcelona, Spain

KEYWORDS. *Active site, esterase, metagenomics, modeling, PELE, rational design.*

ABSTRACT: Effects of altering the properties of an active site in an enzymatic homogeneous catalyst have been extensively reported. However the possibility of increasing the number of such sites, as commonly done in heterogeneous catalytic materials, remains unexplored, particularly because those have to accommodate appropriate residues in specific configurations. This possibility was investigated by using a serine ester hydrolase as the target enzyme. By using the Protein Energy Landscape Exploration software, which maps ligand diffusion and binding, we found a potential binding pocket capable of holding an extra catalytic triad and oxyanion hole contacts. By introducing two mutations, this binding pocket became a catalytic site. Its substrate specificity, substrate preference and catalytic activity were different to those of the native site of the wild type ester hydrolase and other hydrolases, due to the differences in the active site architecture. Converting the binding pocket into an extra catalytic active site was proven a successful approach to create a serine ester hydrolase with two functional reactive groups. Our results illustrate the accuracy and predictive nature of modern modeling techniques, opening novel catalytic opportunities coming from the presence of different catalytic environments in single enzymes.

INTRODUCTION

A number of studies have shown the utility of placing multiple functional groups randomly or uniformly on the surface of solid heterogeneous catalysts to enhance activity and selectivity.¹ Functional groups are positioned in those catalysts at varying distances and spatial arrangements. However, these advances have not been as well developed for enzymatic catalysts. Thus, with few exceptions,²⁻⁴ all enzymes contain one active site. Although introducing single active sites into non-catalytic proteins has been successfully achieved,⁵⁻¹⁶ no evidence has been gathered that demonstrates that introducing a second or more active sites into an enzyme is feasible. If one could use molecular modeling to discover extra binding pockets and convert them into catalytic sites, then one would facilitate designing enzymes with multiple reactive groups. This would promote the competitiveness and catalytic opportunities of enzymes coming from different catalytic environments, and be a considerable step forward in the field of *de novo* enzyme design.

We tested this hypothesis by using a serine ester hydrolase from the structural superfamily of α/β -hydrolases as a model enzyme. The activity of serine ester hydrolases

relies mainly on a catalytic triad usually formed by Ser, Asp/Glu and His.¹⁷ The initial attack of an ester is mediated by the Ser residue that acts as nucleophile through which a covalent intermediate is formed in collaboration with two other residues (Asp and His) that help activate the nucleophile by forming a charge-relay network. At basic pH the His residues are δ -protonated, and the Asp residues are deprotonated, resulting in the formation of a His-Ser and His-Asp hydrogen-bonding network. The stabilization of the intermediate is increased by so-called oxyanion hole contacts with nitrogen atoms of adjacent Gly residues.¹⁷

This enzyme class was selected for a number of reasons. First, it is widely distributed in the environment, it has important physiological functions, it includes hydrolases which are among the most important industrial biocatalysts, and extensive biochemical knowledge has been accumulated.¹⁸ Second, a number of studies have disclosed the utility of computational tools to introduce active sites supporting ester hydrolysis in a number of non-catalytic proteins. A first attempt used RosettaMatch to search for catalytically inert theozyme scaffold proteins that could accommodate a model catalytic triad (Cys-His-

Glu/Asp) and oxyanion holes reassembling cysteine-like geometries.¹¹ Authors identified 4 protein scaffolds with minimalistic catalytic schema consisting of a Cys nucleophile, a nearby His and a backbone NH group, which were capable of hydrolyzing the model esters coumarin-(2-phenyl)-propanoate and a *p*-nitrophenyl-(2-phenyl)-propanoate (pN2PP) (k_2/K_S up to 309 M⁻¹s⁻¹). More recently, Moroz and colleagues took advantage of existing enzymatic folds to explore whether calmodulin (CaM) could assume esterase activity against pN2PP.¹⁶ Docking was used to determine whether pN2PP accommodates into a hydrophobic cavity of CaM. Residues facing the substrate were then identified and further mutated to His, a residue that, per se, has been shown to catalyze hydrolysis of pN2PP, albeit with a modest activity. A mutant with a single Met₁₄₄His substitution was capable of hydrolyzing pN2PP (k_{cat}/K_M up to 4800 M⁻¹min⁻¹). In another approach, the hydrolysis of *p*-nitrophenyl acetate (pNPA) was achieved by incorporating mutations favoring metal binding, with the metal being the catalytic component. By using this approach, Zn²⁺ and/or Hg²⁺ ions were introduced into the Rab-4-binding domain,¹⁰ the TRI family of peptides,¹² and short amyloid-forming peptides.¹⁴ The resulting metallo-proteins were capable of pNPA hydrolysis (k_{cat}/K_M up to 630, 23.3 and 18 M⁻¹s⁻¹). The question of where to put the functional group has been found to be critical for catalytic performance. For example, it has been demonstrated that the His orientation modulates local metal orientation, which in turn has functional consequences.^{19,20}

It remains to be established whether the computational approaches described above can be implemented to introduce multiple active sites supporting ester hydrolysis, not only in non-catalytic proteins but also in ester hydrolases already containing a native functional group. The novelty of this study relies on using the Protein Energy Landscape Exploration (PELE) software. PELE offers one of the best modeling alternatives to map protein-ligand dynamics and induced fit.^{21–25} It allows for a complete protein surface exploration, locating binding site pockets in only few hours of a moderate computing cluster (~32 computing cores).^{26,27} Thus, these technological developments are ideally suited to locate, without previous knowledge, potential binding sites that could be converted into active sites. Finding these pockets typically requires significant enzyme reorganization, both at the level of conformational sampling and induced fit.²⁷ Thus, similar analysis cannot be performed with simple docking techniques. Effective alternative techniques are mostly limited to molecular dynamic simulations,²⁸ although at a significantly higher computational cost.

The present study adds important insights and empirical and computational data proving, for first time to the best of our knowledge, that introducing extra catalytic reactive groups into a serine ester hydrolase is plausible. We would like to highlight that our approach is based on locating, by PELE software, extra binding pockets and converting them into catalytic active sites. We were interested not only in proving that introducing extra function-

al groups into a catalytic serine ester hydrolase is plausible, but also in examining the catalytic potential of the resulting variant compared to the wild type enzyme and other native ester hydrolases. In this line, we have unambiguously confirmed that the extra catalytic site is not only competitive with those of other native ester hydrolases, but also could confer catalytic changes when introduced in the wild type ester hydrolase already containing a native active site.

EXPERIMENTAL AND COMPUTATIONAL METHODS

Chemicals and oligonucleotides. All chemicals used for enzymatic tests were of the purest grade available and were purchased from Sigma Chemical Co. (St Louis, MO, USA), Alfa Aesar (Karlsruhe, Germany) or Santa Cruz Biotechnology (Heidelberg, Germany). The oligonucleotides used for DNA amplification were synthesized by Sigma Genosys Ltd. (Pampisford, Cambs, UK).

EH_{1A} protein source and crystal structure. The isolation of the enzyme EH_{1A} used in the present study was reported previously.²⁹ The enzyme is available in the expression vector pET46 Ek/LIC plasmid and *Escherichia coli* BL21 as a host,^{29,30} which was the source of the enzyme for the present study. The crystal structure of EH_{1A} protein was recently solved (PDB code: 5JD4) and X-ray diffraction data collection and refinement statistics are available.³⁰

Site directed mutagenesis. Mutagenic PCR was developed using the Quick Change Lightning Multi Site-Directed Mutagenesis kit (Agilent Technologies, Cheadle, UK), following manufacturer instructions. Briefly, 50 ng of pET46 Ek/LIC plasmid containing wild type EH_{1A} DNA insert was mixed with a master mix containing 2.5 µl of 10x Multireaction buffer, 0.5 µl of Quick solution, 1 µl of dNTP mix, 1 µl of Multy enzyme blend and 100 ng of each primer. Distilled water was added to a final volume of 25 µl. PCR conditions were as follow: 2 minutes at 95°C, 30 cycles of 20 seconds at 95°C, 30 seconds at 55°C, 195 seconds at 65°C, and one cycle of 5 minutes at 65°C. The resulting variant plasmids were transferred into *E. coli* BL21 and selected on the Luria Bertani (LB) agar supplemented with 50 µg ml⁻¹ ampicillin. To obtain the variant EH_{1B}, the following primers were used: Ser161Ala Fwd (GTG GGC GGC GAT GCG GCG GGC GGC G), Glu25Asp Fwd (CGG CCC CGG CTG GAT ACC CTG CCG CAT GC) and Leu214His Fwd (TTC CTC AGC AAG GCG CAC ATG GAC TGG TTC TGG G). To obtain the variant EH_{1AB}, only the primers Glu25Asp Fwd and Ile214His Fwd were used.

To obtain variants containing Ser₂₁₁Ala, Ser₁₆₁Ala, Asp₂₅Gln, Asp₂₅₆Gln, His₂₁₄Phe and His₂₈₆Phe mutations, individually or in combination, the pET46 Ek/LIC plasmid containing EH_{1A}, EH_{1B}, or EH_{1AB} DNA inserts and the primers Ser₂₁₁Ala Fwd (GCC GAA GGC TAC TTC CTC GCC AAG GCG CAC ATG GAC TGG), Ser₁₆₁Ala Fwd (GTG GGC GGC GAT GCG GCG GGC GGC G), Asp₂₅Gln Fwd (CGG CCC CGG CTG CAG ACC CTG CCG CAT GC), Asp₂₅₆Gln Fwd (ACC GCC GGC TAC CAA CCG CTG CGC GAC G), His₂₁₄Phe Fwd (TTC CTC AGC AAG GCG TTC

ATG GAC TGG TTC TGG G) and His286Phe Fwd (T CCC GGC ACC ATC TTC GGC TTC TTC TCG), were used. Mutagenic PCR conditions were as above.

Gene expression and protein purification. Protein expression and purification of wild type and mutants were performed as previously described with slight modifications.^{29,30} Briefly, selected *E. coli* clones that expressed each protein, His-tagged at the N-terminus, were grown at 37 °C on solid LB agar medium supplemented with 50 µg ml⁻¹ ampicillin, and one colony was picked and used to inoculate 10 ml of LB broth plus antibiotic in a 0.25-l flask. The cultures were then incubated at 37 °C and 200 rpm overnight. Afterward, 10 ml of this culture was used to inoculate 0.5 liter of LB medium, which was then incubated to an OD_{600nm} to approximately 0.7 (ranging from 0.55 to 0.75) at 37 °C. Protein expression was induced by adding isopropyl β-D-1-thiogalactopyranoside to a final concentration of approx. 1 mM, followed by incubation for 16 h at 16 °C. The cells were harvested by centrifugation at 5000 × *g* for 15 min to yield a pellet of 2–3 g l⁻¹ pellet (wet weight). The wet cell pellet was frozen at -86 °C overnight, thawed and resuspended in 15 ml of 40 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 7.0. Lysozyme Bioprocessing Reagent (Novagen, Darmstadt, Germany) was added (4 µl g⁻¹ wet cells) and incubated for 60 min on ice with rotating mixing. The cell suspension was sonicated using a pin Sonicator® 3000 (Misonix, New Highway Farmingdale, NY, USA) for a total time of 5 min (10 W) on ice and centrifuged at 15000 × *g* for 15 min at 4 °C, and the supernatant was retained.

The His-tagged proteins, native and engineered variants, were purified at 4 °C after binding to a Ni-NTA His-Bind resin (Sigma Chemical Co. (St Louis, MO, USA)), followed by ultra-filtration through low-adsorption hydrophilic 10000 nominal molecular weight limit cutoff membranes (regenerated cellulose, Amicon) to concentrate the protein solution. An extensive dialysis of protein solutions against 40 mM HEPES buffer (pH 7.0) was then performed using Pur-A-Lyzer™ Maxi 1200 dialysis kit ((Sigma Chemical Co. (St Louis, MO, USA)), as follows. Five ml concentrated protein solution was dialyzed against 2 L buffer during 1 hour at room temperature, after which the buffer was changed by other 2 L buffer and maintained 1 hour more. Then, the buffer was changed and the dialysis was kept overnight at 4 °C. The dialyzed protein solution was recovered and concentrated as before. The concentrated protein solution (10 mg ml⁻¹) was then subjected to size-exclusion chromatography by a fast protein liquid chromatography (FPLC) equipment (LCC-500CI, Amersham Bioscience, Barcelona, Spain). The protein sample was loaded onto the FPLC coupled with a Superdex 75 size exclusion column pre-equilibrated with buffer HEPES buffer (pH 7.0). The proteins were eluted with the same buffer at a flow rate of 1 ml min⁻¹. Fractions with hydrolytic activity were pooled, concentrated, and dialyzed against HEPES buffer (pH 7.0), as before. Throughout the purification protocol the fractions were analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) on 12 % gels, in a Mini

PROTEAN electrophoresis system (Bio-Rad),³¹ in which the proteins were stained with Coomassie brilliant blue (Problue Safe, National Diagnostics, GA, USA), and for hydrolytic activity using *p*-nitrophenyl propionate (pNPP) as described below. The protein concentration was determined according to Bradford with bovine serum albumin as the standard.³²

Purity was assessed as >99% by SDS-PAGE, Matrix-Assisted Laser Desorption/Ionization-Time-Of-Flight/Time-Of-Flight (MALDI-TOF/TOF) and other complementary techniques (see Supporting Results).

Ester bond hydrolysis activity assessment: substrate profiling tests with 96 esters. Hydrolytic activity was assayed at 550 nm using structurally diverse esters in 384-well plates as previously described.^{30,33} All chemicals used were of the purest grade available. Briefly, before an assay, a concentrated ester stock solution was prepared in a 96-well plate by dissolving each of the 96 esters at a concentration of 25 mg ml⁻¹ in acetonitrile or dimethyl sulfoxide (DMSO), depending on its solubility. Stock solutions were prepared immediately prior to use and maintained in a 96-deep-well plate at 4 °C.

The assays were conducted according to the following steps. First, a 384-well plate (Molecular Devices, LLC, CA, USA) was filled with 20 µl of 5 mM N-(2-hydroxyethyl)piperazine-N'-(3-propanesulfonic acid (EPPS) buffer, pH 8.0, using a QFill3 microplate filler (Molecular Devices, LLC, CA, USA). Second, 2 µl of each ester stock solution was added to each well using a PRIMADIAG liquid-handling robot (EYOWN TECHNOLOGIES SL, Madrid, Spain). Each of the 96 esters was dispensed in four replicates in each 384-plate. After adding the esters, the 384-well plate was filled with 20 µl of 5 mM EPPS buffer, pH 8.0, containing 0.912 mM Phenol Red (used as a pH indicator) using a QFill3 microplate filler. The final ester concentration in each well was 1.14 mg ml⁻¹, and the final concentration of Phenol Red was 0.45 mM. A total of 2 µl of protein solution (from a 1 (for EH_{1A}/EH_{1AB}) or 16 (for EH_{1B}) mg ml⁻¹ stock solution in 40 mM HEPES buffer pH 7.0) was immediately added to each well using an Eppendorf Repeater M4 pipette (Eppendorf, Hamburg, Germany) or a PRIMADIAG liquid-handling robot. Accordingly, the total reaction volume was 44 µl, with 4.5% (v/v) acetonitrile or DMSO in the reaction mixture. After incubation at 30 °C and 150 rpm in a Synergy HT Multi-Mode Microplate Reader, ester hydrolysis was measured spectrophotometrically in continuous mode at 550 nm for a total time of 24 h. One unit (U) of enzyme activity was defined as the number of enzyme required to transform 1 µmol of substrate in 1 min under the assay conditions using the reported extinction coefficient ($\epsilon_{\text{Phenol red}}$ at 550 nm = 8450 M⁻¹ cm⁻¹).²⁹ All values were corrected for non-enzymatic transformation.

Kinetic measurements. For determination of kinetic parameters, these were calculated by simple Michaelis-Menten kinetics.

Kinetics experiments for pNPP were initiated by the addition of a stock solution (100 mM in acetonitrile) of

*p*NPP to 195 μ l of HEPES buffer (pH 7.0) containing EH_{1A}, EH_{1B} or EH_{1AB} in 96-well microtiter plates. Kinetics experiments for phenyl propionate were initiated by the addition of a stock solution (100 mM in acetonitrile) of phenyl propionate to 40 μ l of 5 mM EPPS buffer pH 8.0 with 0.45 mM Phenol Red (used as a pH indicator) containing EH_{1A}, EH_{1B} or EH_{1AB} in 96-well microtiter plates. For K_m determinations, the amount of protein was 0.1, 16 and 1.4 μ g for EH_{1A}, EH_{1B} and EH_{1AB}, respectively, or 100 μ g for knockout mutants. For k_{cat} determinations, the amount of *p*NPP and phenyl propionate was 1 and 6 mM, respectively.

In all cases, reactions were followed at 410 nm (for *p*NPP hydrolysis) or 550 nm (for phenyl propionate hydrolysis) by UV-vis spectrophotometry in a Synergy HT Multi-Mode Microplate Reader. Initial rates were determined from linear fits of the absorbance versus time (<10% conversion) corrected for the rate of un-catalyzed hydrolysis. In all cases, one unit (U) of enzyme activity was defined as the number of enzyme required to transform 1 μ mol substrate in 1 min under the assay conditions using the reported extinction coefficient (ϵ_{pNPP} at 410 nm = 15200 M⁻¹ cm⁻¹; $\epsilon_{phenol\ red}$ at 550 nm = 8450 M⁻¹ cm⁻¹). In all cases, reactions (performed in triplicate) were maintained at 30°C.

pH optima determination. Britton and Robinson buffer (50 mM; pH 4.0–9.5) and *p*NPP (1 mM) was used for determining optimal pH of the enzyme variants. A total of 2 μ l of a stock solution (100 mM in acetonitrile) of *p*NPP were added to 195 μ l of HEPES buffer (pH 7.0) containing EH_{1A} (0.1 μ g), EH_{1B} (16 μ g) or EH_{1AB} (1.4 μ g) in 96-well microtiter plates. The *p*NPP hydrolysis was monitored as above but at 346 nm (ϵ_{pNPP} at 346 nm = 4800 M⁻¹ cm⁻¹ regardless of solution pH), in triplicates.

Protein Energy Landscape Exploration (PELE) simulations. The initial structure was taken from the coordinates of the EH_{1A} crystal structure (PDB code: 5JD4).³⁰ The protonation state of titratable residues was estimated with the Protein Preparation Wizard (PROPKA)³⁴ and the H++ server (<http://biophysics.cs.vt.edu/H++>) followed by visible inspection. At pH 8 (the pH at which the activity assays were performed), the catalytic triad histidine residues were δ -protonated, and the catalytic triad aspartic acid residues were deprotonated, resulting in the formation of a histidine-serine and histidine-aspartic hydrogen-bonding network. The glyceryl tri-propionate structure was fully optimized with Jaguar²⁷ in an implicit solvent, and the electrostatic potential charges were computed with the density functional Mo6 at the 6-31G* level of theory; ligand parameters were extracted from these for the classic simulations.

We used Protein Energy Landscape Exploration (PELE) software to sample the binding mode of glyceryl tri-propionate with EH_{1A}.²¹ PELE is a Monte Carlo algorithm composed of a sequence of perturbation, relaxation, and Metropolis acceptance tests. In the first step, the ligand is subjected to random rotations and translations, while the protein is perturbed based on the anisotropic network

model (ANM).²⁷ The maximum allowed translation for the ligand perturbation was 1.5 Å, and the maximum rotation was 20°. During the protein perturbation, all atoms were displaced by a maximum of 0.5 Å by moving the α -carbons following a random linear combination of the 6 lowest eigenvectors obtained in the ANM model. The relaxation step included the repositioning of all amino acid side chains within 6 Å of the ligand and the 5 side chains with the highest energy increase along the previous ANM step. The relaxation stage ended with a truncated Newton minimization using the OPLS all-atom force field and an implicit surface-generalized Born continuum solvent.²⁶ The new proposed minima were then accepted or rejected based on a Metropolis test. The substrate binding plots contained all accepted conformations for three 12-h simulations using 200 processors.

Molecular dynamics. 250 ns of molecular dynamics (MD) simulation with DESMOND³⁵ were performed to assure the enzymatic stability. After appropriate preparation of the system, as explained before, an orthorhombic water box with a minimum distance of 10 Å was introduced. The systems were then neutralized and 150 mM NaCl added. Equilibration using the default protocol was performed followed by 20 ns NPT simulation at 300 K and 1 atm with the OPLS-2005 force field. The temperature was regulated with the Nosé-Hoover chain thermostat while the pressure was controlled by the Martyna-artynae was barostat with isotropic coupling and a relaxation time of 2.0 ps.

Peptide mass fingerprinting by Matrix-Assisted Laser Desorption/Ionization-Time-Of-Flight/Time-Of-Flight (MALDI-TOF/TOF). Before MALDI-TOF/TOF analysis in-solution, protein digestion was performed.³⁶ Briefly, 20 μ g of protein samples (5 mg ml⁻¹ in 40 mM HEPES buffer pH 7.0) were diluted and denatured in 20 μ l 7 M urea/2 M thiourea/100 mM triethylammonium bicarbonate (TEAB), pH 7.5, reduced with 2 μ l of 50 mM tris(2-carboxyethyl) phosphine (TCEP, AB SCIEX), pH 8.0, at 37°C for 60 min, followed by addition of 2 μ l of 200 mM cysteine-blocking reagent (methyl methanethiosulfonate (MMTS); Pierce, Appleton, WI, USA) for 10 min at room temperature. Samples were diluted up to 120 μ l to reduce guanidine concentration with 50 mM TEAB. Digestions were performed using sequence grade-modified trypsin (Promega, Alcobendas, Spain) to each sample in a ratio 1/20 (w/w), which were then incubated at 37°C overnight on a shaker. Sample digestions were evaporated to dryness and were cleaned-up/desalted using Stage-Tips with Empore 3M C18 disks (Sigma Chemical Co.; St Louis, MO, USA). The tryptic eluted peptides were dried by speed-vacuum centrifugation and resuspended in 4 μ l of MALDI solution (30% acetonitrile/15% isopropanol/0.5% trifluoroacetic acid). A 0.8 μ l aliquot of each peptide mixture was deposited onto a 384-well OptiTOF™ Plate (SCIEX, Foster City, CA) and allowed to dry at room temperature. A 0.8 μ l aliquot of matrix solution (3 mg ml⁻¹ α -cyano-4-hydroxycinnamic acid in MALDI solution) was then deposited onto dried digest and allowed to dry at room temperature.

For MALDI-TOF/TOF analysis, samples were automatically acquired in an ABI 4800 MALDI TOF/TOF mass spectrometer (SCIEX, Foster City, CA) in positive ion reflector mode (the ion acceleration voltage was 25 kV to MS acquisition and 2 kV to MSMS) and the obtained spectra were stored into the ABI 4000 Series Explorer Spot Set Manager. PMF and MSMS fragment ion spectra were smoothed and corrected to zero baseline using routines embedded in ABI 4000 Series Explorer Software v3.6. Each PMF spectrum was internally calibrated with the mass signals of trypsin autolysis ions to reach a typical mass measurement accuracy of <25 ppm. Known trypsin and keratin mass signals, as well as potential sodium and potassium adducts (+21 Da and +39 Da) were removed from the peak list. To submit the combined PMF and MS/MS data to MASCOT software v.2.6.0 (Matrix Science, London, UK), GPS Explorer v4.9 was used, searching in a custom protein database with the sequences encoding EH_{1A}, EH_{1B} and EH_{1AB}. The following search parameters were used: enzyme, trypsin; allowed missed cleavages, 1; methylthiolation cysteine as fixed modification by the treatment with MMTS; variable modifications, oxidation of methionine; mass tolerance for precursors was set to ± 50 ppm and for MS/MS fragment ions to ± 0.3 Da. The confidence interval for protein identification was set to $\geq 95\%$ ($p < 0.05$) and only peptides with an individual ion score above the identity threshold were considered correctly identified.

Estimation of molecular mass by MALDI-TOF/TOF. Protein samples were diluted at 1:1 ratio (v/v) with matrix solution (50% saturated sinapinic acid in 70% aqueous acetonitrile and 0.1% trifluoroacetic acid). A 1.0 μ l aliquot of this mixture was manually deposited onto a 386-well OptiTOF™ Plate (ABSciex, Framingham, MA, USA) and allowed to dry at room temperature. For MALDI-TOF/TOF analysis, samples were automatically acquired in an ABI 4800 MALDI TOF/TOF mass spectrometer (SCIEX, Foster City, CA) in positive ion linear mode (the ion acceleration voltage was 25 kV for MS acquisition). The detection mass range was set between 1,500 and 80,000 m/z .

RESULTS AND DISCUSSION

Model ester hydrolase for this study. We chose a serine ester hydrolase, herein referred to as EH_{1A}, with a typical α/β hydrolase fold as a model. In a recent study, it was identified as the most promiscuous ester hydrolase among a total of 147 when tested with a set of 96 chemically and structurally different esters (Supporting Table S1).³⁰ EH_{1A} was isolated from the metagenomic DNA of microbial communities inhabiting a karstic lake,²⁹ and its structure was recently solved (PDB code 5JD4).³⁰ The active site is formed by a catalytic triad formed by Ser161, Asp256 and His286 (Figure 1) and an oxyanion hole formed by Gly88, Gly89 and Gly90.³⁰ This active site supports the hydrolysis of a broad range of 72 esters (Figure 2, inset), with vinyl butyrate and phenyl propionate serving as best substrates (Figure 2).

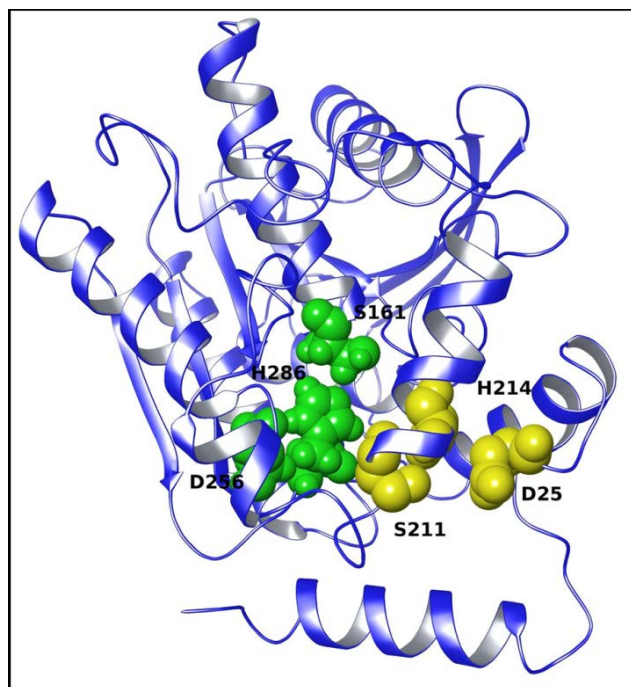


Figure 1. The relative position of the two active sites, with their catalytic triads. Original EH_{1A} catalytic triad and oxyanion hole are shown in green and the designed EH_{1B} catalytic triad and oxyanion hole in yellow.

PELE simulations for locating an extra binding pocket. Using the PELE software, which allows mapping ligand diffusion and binding,^{21–25} we performed an exploration of EH_{1A} with glyceryl tri-propionate, a ligand with high activity in multiple ester hydrolases, including EH_{1A} ($\sim 62000 \pm 7400$ U g⁻¹; Figure 2), and appropriately sized for binding into defined cavities. The rationale was to identify (alternative) potential binding sites to accommodate a new active site (Figure 3). PELE simulations revealed a potential second binding site located ~ 13 Å from the native catalytic position at Ser161 (Figure 4a). This second site (Figure 4b) already contains a serine residue (Ser211). Thus, we computationally designed additional mutations, adding Asp and His residues to build a proper catalytic triad, taking special care of distances between the residues and substrate accommodation. PELE results for the Glu25Asp and Leu214His double mutant revealed good enzyme-substrate interaction energies (Supporting Figure S1) and a suitable catalytic position for the glyceryl tri-propionate substrate. Figure 4 summarizes the catalytic triad environment of the newly introduced active site compared to the wild type. In addition, our results show that residues Gly207, Tyr208 and Phe209 act as a potential oxyanion hole, a key element in ester hydrolase catalysis (Figure 4b). Moreover, extensive MD simulations indicate proper stabilization of the double mutant (Supporting Figure S2).

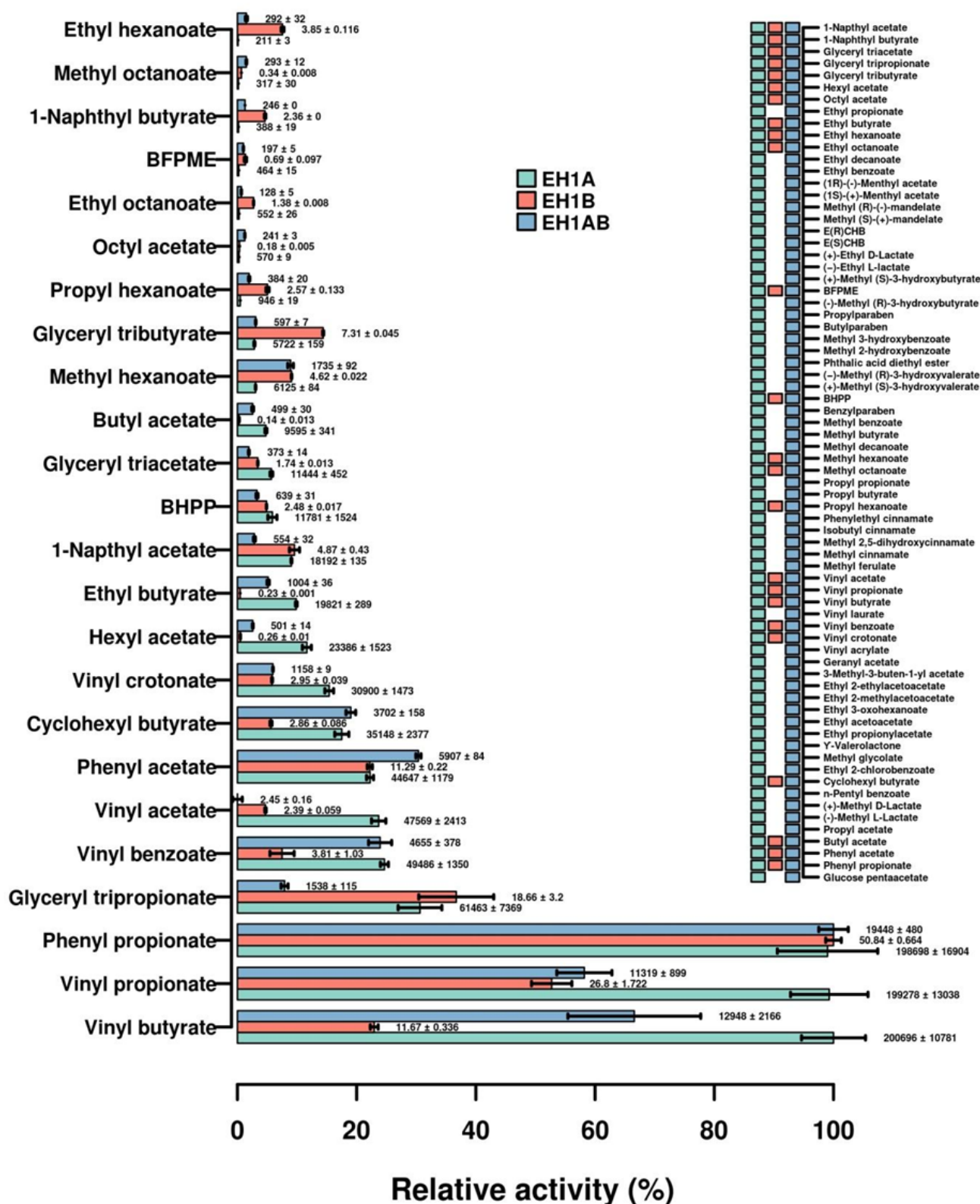


Figure 2. Substrate spectra of EH1 variants. The 72 out of 96 esters for which activity was detected are listed on the right side panel, with indication of substrates being converted by each of the 3 variants. The ID code representing each EH variant is color coded. On the left main panel the relative activity (%) for each ester referred to the best substrate is indicated, with specific activity (U g^{-1}) and standard deviation (triplicates) indicated in the bars. Abbreviations as follows: E(R/S)CHB: ethyl (R/S)-(+)-4-chloro-3-hydroxybutyrate; BFPME: benzoic acid, 4-formyl-, phenylmethyl ester; BHPP: benzyl (R)-(+)-2-hydroxy-3-phenylpropionate. The activity protocol established and used to identify the esters hydrolyzed by each variant was based on a 550-nm follow-up pH indicator assay at pH 8.0 and 30 °C (see Experimental section).

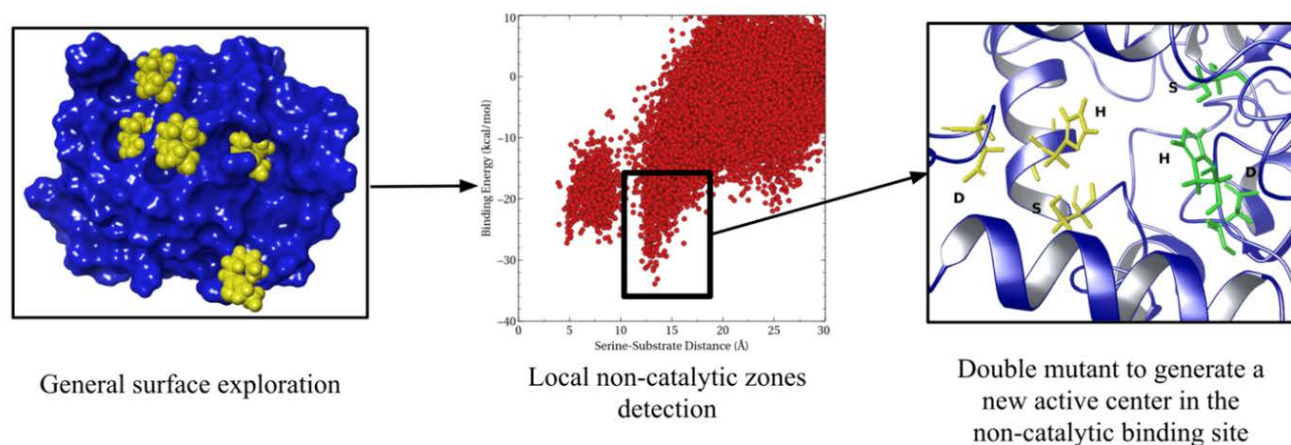


Figure 3. Molecular modeling protocol used for adding a second catalytic active site: i) Complete surface exploration with the probe ligand (glyceryl tri-propionate); ii) Alternative binding site location; iii) Substrate placement simulations (local refinement on the new site) with mutations.

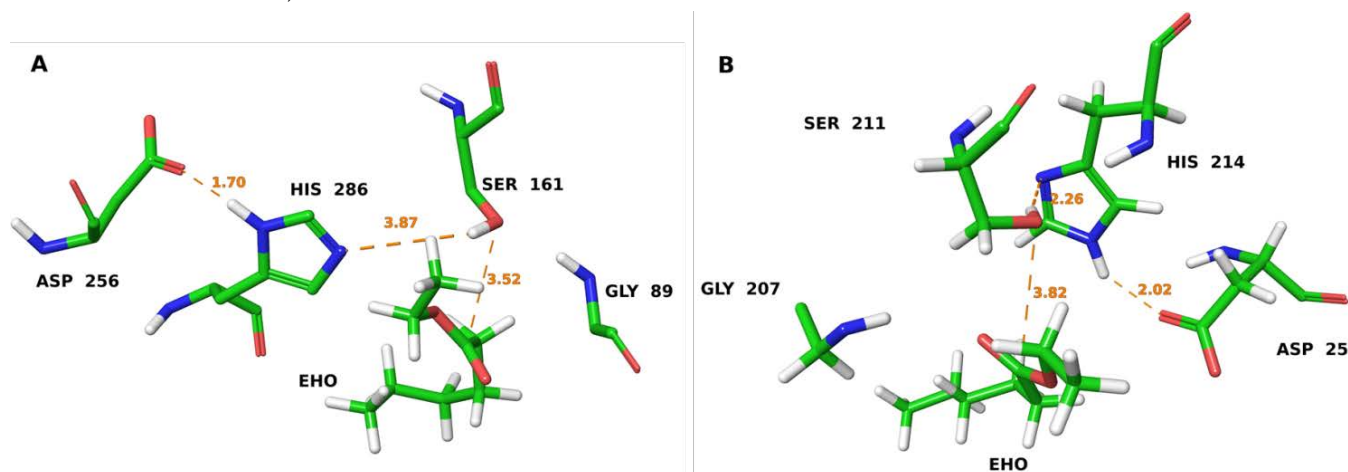


Figure 4. Catalytic position adopted by ethyl hexanoate (EHO) in the wild type active site (a) and in the Glu25Asp and Leu214His double mutant (b). Along with the substrate, the catalytic triad and the oxyanion hole residue are shown; main catalytic distances are underlined in angstroms.

Converting the binding pocket into a functional catalytic site by site-directed mutagenesis. By using site-directed mutagenesis and with the above considerations, an enzyme variant with the new presumptive active site and an inactive original site was designed to determine whether the extra active site was, per se, functional. We introduced Glu25Asp, Leu214His and Ser161Ala substitutions, so that this variant, named EH_{1B}, would presumably employ a new catalytic triad (Ser211, Asp25 and His214) with Ser211 as the nucleophile and a new oxyanion hole (Gly207, Tyr208 and Phe209). The corresponding gene was cloned, and the His₆-tagged protein was expressed and purified by Ni-NTA affinity and size-exclusion chromatography (Supporting Figure S3 and Table S2). A number of control experiments were performed to ensure the purity and identity of the EH_{1B} protein (Supporting Figures S3-S8). These analyses were also performed with purified EH_{1A} and other variant (see below). With the sensitivity allowing MALDI-TOF/TOF and other complementary techniques (see Supporting Results), we can conclude that there is no contamination

and that the purity of the EH_{1B} (as well as EH_{1A} and further variant EH_{1AB}) was higher than 99%. The introduced mutations were confirmed in all cases by sequencing of the corresponding gene and by peptide mass fingerprinting by MALDI-TOF/TOF of purified proteins (for details see Supporting Results, and Figures S5 and S6).

Once the purity and nature of EH_{1B} was confirmed, its hydrolytic activity was evaluated against the set of 96 esters as for the wild-type ester hydrolase (see Experimental section). We found that EH_{1B} with the new active site showed not only measurable activity but also catalyzed reactions with 24 different substrates (Figure 2), including cyclohexyl butyrate and the large aromatic ester benzoic acid, 4-formyl-, phenylmethyl ester, which are rarely hydrolyzed by many native ester hydrolases (Figure 5; esters ID nr. 59 and 67).³⁰ Note that in EH_{1B} the original active site (Ser161) was silenced by mutating it by Ala, mutation per se found to completely abolish the activity of the original active site (Supporting Figure S9D). Thus, the capacity of EH_{1B} to convert these 24 esters could be unambiguously assigned to the new active site.

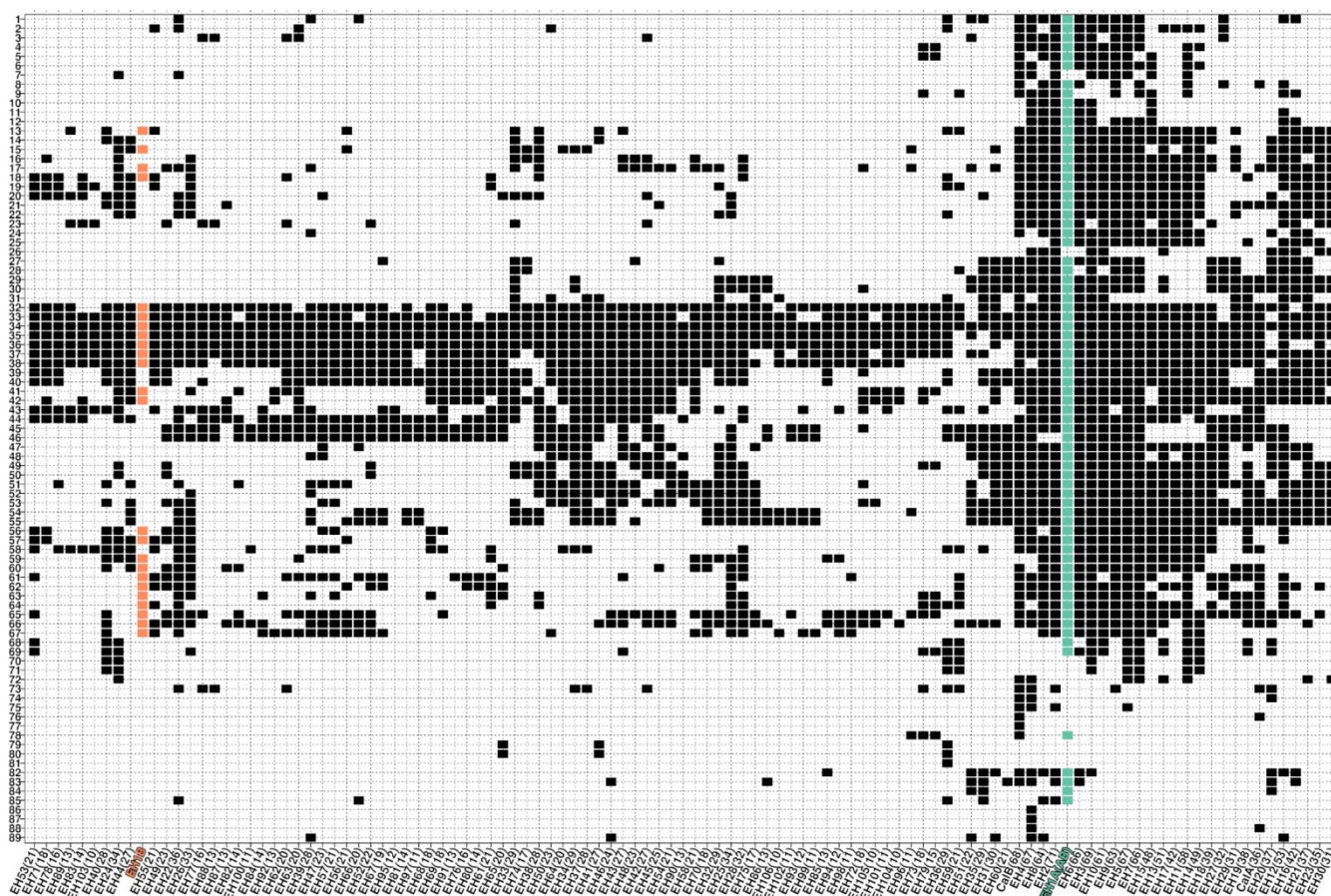


Figure 5. Clustering of the substrate range of EH1 variants within them and in comparison with those of other serine ester hydrolases. A total of 107 ester hydrolases, including the commercial preparations CalA and CalB from *Pseudozyma aphidis*, formerly *Candida antarctica*, are included. The data corresponding to each EH variant is color coded. This figure is created from data previously reported³⁰ and data herein obtained for EH1 variants (see Experimental section). The list of the esters tested is shown on the left side (full name of the esters from nr. 1 to 89 is given in Supporting Table S4). The ID code representing each ester-hydrolase is given at the bottom. Each hydrolase is named based on the code ‘EH’, which means Ester Hydrolase, followed by an arbitrary number. The number in brackets indicates the number of esters hydrolyzed by each enzyme. The Figure was created with the R language console, as described previously,³⁰ using a binomial table with information about the activity/inactivity (1/o) of the analyzed enzymes against the 96 substrates as a starting point.

To evaluate the catalytic potential of the newly introduced active site, we compared the number of esters being hydrolyzed with that of the wild-type protein (EH_{1A}), 144 naturally occurring ester hydrolases and two commercial preparations tested against the same set of 96 esters.³⁰

As shown in Figure 6, we observed that EH_{1B} had a restricted substrate spectrum (24 esters) compared to EH_{1A} (72 esters), the most promiscuous ester hydrolase among those examined, but it was broader than that of the other hundred native ester hydrolases. The ester hydrolase with the new active site designed herein would fall into the category of ester hydrolases with moderate substrate promiscuity, thus suggesting the hydrolytic potential of the newly introduced active site in terms of the substrate spectrum.

We further evaluated whether EH_{1B} showed different substrate preferences compared to EH_{1A}. We considered only those 24 substrates converted by both variants (Figure 2) and calculated the relative activity for each. As shown in Figure 2, distinct profiles were observed that highlight the difference in binding capacity for each of the catalytic environments. As an example, the new active site (EH_{1B}) showed a higher relative preference for the hydrolysis of methyl-, propyl- and ethyl-hexanoate compared to the native site (EH_{1A}). The difference in substrate binding and preference between the new (EH_{1B}) and wild type (EH_{1A}) active sites was also noticeably when comparing their capacity to hydrolyze phenyl propionate and pNPP (Supporting Figure S9 and Table S3). Thus, in term of k_{cat} EH_{1B} preferred phenyl propionate over pNPP (5.7-fold), whereas EH_{1A} slightly preferred pNPP (1.4-fold). A

local refinement simulation with PELE revealed that the changes in preference for ethyl-hexanoate correlate with substrate placement. In particular, the difference in number of catalytic events in both sites, which define the poses where the substrate is readily placed for catalysis, is significantly reduced for ethyl-hexanoate, when compared to one of the substrates with maximum activity, phenyl propionate (Figure 7). Thus, differences in substrate profile may be due to distinct active site architectures (Figures 1 and 4). Additionally, there are different active site effective volumes (the active site cavity volume corrected by the relative solvent accessible surface area (SASA) of the catalytic triad) for the EH_{1B} active site (82.2 Å³) and the EH_{1A} active site (167.7 Å³). This volume has been recently correlated with changes in substrate promiscuity in ester hydrolases.³⁰

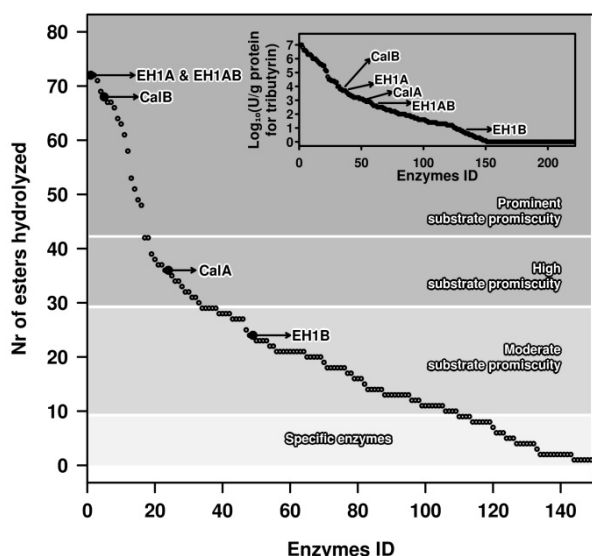


Figure 6. Comparison of the substrate range and catalytic performance of EH₁ variants and other reported ester hydrolases. This figure is created from data previously reported³⁰ and data herein reported for EH₁ variants (see Figures 2 and 5). The main figure represents the number of substrates hydrolyzed by each EH₁ variant compared to that of other 146 serine ester hydrolases that included the commercial preparations CalA and CalB.³⁰ The inset represents the specific activity against glyceryl tri-butyrate of the EH₁ variants and other 225 reported ester hydrolases (that included the 146 previously mentioned³⁰ and 79 reported elsewhere).

Finally, we compared the hydrolytic rate of EH_{1B} with that of EH_{1A} and 225 previously reported ester hydrolases. We chose glyceryl tri-butyrate as a representative ester substrate because its specific activity values are commonly reported. As shown in Figure 2, the activity (7.3 ± 0.1 U g⁻¹) might seem moderate compared to that of EH_{1A} (5722 ± 159 U g⁻¹). This was also proven by calculating the k_{cat}/K_M for phenyl propionate and *p*NPP, which was also found as a substrate (see Supporting Figure S9 and Table

S3). Based on the modeled structure (Figure 4), the designed side chains were nearby but not quite in the typical orientation for the catalytic triad, which may explain why catalysis was less effective for EH_{1B} compared to EH_{1A}. It is well known that catalytic triads in serine ester hydrolases require a precise relative orientation of the Ser-Asp-His side chains.¹¹ Whatever the difference with EH_{1A}, we observed that the specific activity of EH_{1B} is comparable to or higher than that of 93 previously reported native ester hydrolases that showed either low (≤ 6.8 U g⁻¹) or no (near zero) capacity to hydrolyze this substrate (about 70 esterases) (Figure 6 inset). The other reported 132 esterases were found far more active for the target substrate. Together, we can conclude that, at least for the hydrolysis of glyceryl tri-butyrate, the designed active site albeit being characterized by low catalytic activity, is as active as those of many natural hydrolases, but still not competitive with those of most esterases.

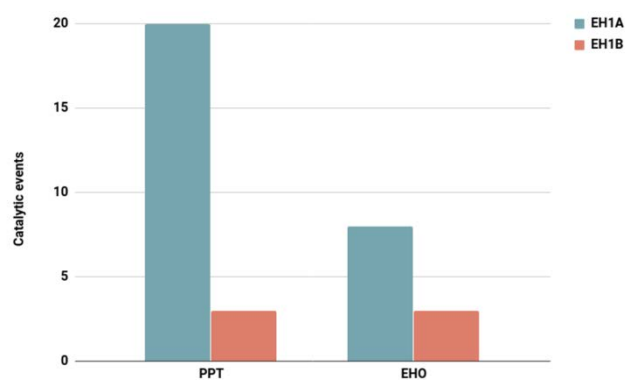


Figure 7. Number of catalytic events found in EH_{1A} and EH_{1B} catalytic sites along PELE simulations for ethyl hexanoate (EHO) and phenyl propionate (PPT). A catalytic event was defined when the main catalytic distances: ligand-Ser, Ser-His, and His-Asp, are lower than 3.5 angstroms. All simulations include the same sampling effort.

We found that single Ser₂₁₁Ala, as well as Asp₂₅Gln or His₂₁₄Phe, substitutions in EH_{1B} completely arrested the activity for all 96 chemically and structurally different esters examined in this study (Supporting Table S2), that included the 24 different esters hydrolyzed by EH_{1B} (see Figure 2), and *p*NPP (see Supporting Results and Figures S9 and S10). This unambiguously confirms that they are the functional groups supporting catalysis in the new active site, which is constituted by a catalytic triad, and not a diad or a histidine. Altogether, these data demonstrated that it is plausible to use molecular modeling to find a potential extra binding pocket in a serine ester hydrolase that can be turned into a catalytic site. This site compares in terms of substrate spectra and catalytic activity with those from many other reported native ester hydrolases. It is difficult to make a statement whether the substrate spectrum of the newly introduced active site is comparable to, or is higher than that in the variants designed *de novo* by computational tools previously de-

scribed.^{10–12,14,16} This is because these variants were only tested with a restricted set of *p*-nitrophenyl esters, substrates that are more easily hydrolyzed compared to the esters tested herein (Figure 2). It is clear though that, at least, EH_{1B} is capable of *p*NPP hydrolysis with a k_{cat}/K_M of 16.4 M⁻¹s⁻¹ (Supporting Figure S9 and Table S3), a value within the range or lower than that reported for other constructs using related substrates (k_{cat}/K_M from 18 to 309 M⁻¹s⁻¹).

Design and characterization of a serine ester hydrolase with two reactive groups. An enzyme variant with the two active sites (the native and the newly identified) was created to prove that an enzyme with two hydrolytic active sites is functional. This mutant, named EH_{1AB}, was created by incorporating the Glu25Asp and Leu214His substitutions into the wild-type sequence. This two-active-site variant would likely employ Ser211 and Ser161 as nucleophiles (Figure 1). The corresponding gene was cloned, and the protein was expressed, purified (Supporting Figure S3 and Table S2) and characterized as described above.

The analysis of the variant EH_{1AB} confirmed its ability to hydrolyze all 72 esters converted by EH_{1A} (Figure 2). Thus, EH_{1AB}, as the native EH_{1A}, could be classified as an ester hydrolase with prominent substrate promiscuity (Figures 5 and 6). Because EH_{1A}, due to its promiscuous behavior, converted all esters hydrolyzed by EH_{1B}, we could not examine the effects of cooperativity for expanding the substrate spectra because of the presence of two different active sites in EH_{1AB}. We further observed that EH_{1AB} does show overall a substrate preference similar to that of EH_{1A} (Figure 2), although in some cases, the potential of the newly introduced active site to increasing the preference for a number of esters, such as methyl-, propyl- and ethylhexanoate, was noticed. Also, EH_{1AB} preferred phenyl propionate over *p*NPP (1.3-fold), whereas EH_{1A} slightly preferred *p*NPP (1.4-fold) (Supporting Table S3). This may be a consequence of a higher preference of the second active site for phenyl propionate over *p*NPP.

The specific activity of EH_{1AB} for all accepted esters was lower than that observed for EH_{1A} but higher than that for EH_{1B} (Figure 2). As example, using glyceryl tri-butyrate, its activity (597 ± 7 U g⁻¹) was 10-fold lower than that of EH_{1A} and 81-fold higher than that of EH_{1B}. These differences were also observed when calculating the k_{cat}/K_M for phenyl propionate and *p*NPP (see Supporting Figure S9 and Table S3). The change in activity with respect EH_{1A} could be explained by the decrease in substrate population for the main active site observed when introducing the mutants (Supporting Figures S1a and b). Visual inspection of the models and energies (Supporting Figure S1b) show a larger connection between the two sites, induced by changes in the hydrophobic packing (Leu24, Phe218, Leu24 and Trp217) after introducing His214. The alpha helix harboring the Leu214His mutation forms one face of the native active site, which slightly perturbs the main site. Whatever the catalytic activity level of the EH_{1AB} variant compared to the native enzyme, its activity

is significantly higher than that observed for 163 naturally occurring ester hydrolases (Figure 6, inset).

The consequences of the presence of two active sites were visible in K_M Michaelis-Menten (Supporting Figure S9 and Table S3) and pH optima (Supporting Figure S11) curves. Thus, the K_M progress curve for EH_{1AB} displays a sigmoidal plot of the initial reaction rate versus substrate concentration (for both *p*NPP and phenyl propionate), rather than the hyperbolic plots observed for EH_{1A} and EH_{1B}. This situation is typically found in allosteric enzymes, where the binding of substrates occurs in several active sites in the same enzyme molecule,³⁷ a situation that may most likely occurs in EH_{1AB}. We also observed that EH_{1AB} displays an optimum pH profile distinct (narrower) to those of EH_{1A} (which showed the broader range) and EH_{1B} (see Supporting Results), which also differ between them (Supporting Figure S11). Subsequent investigation of the reasons that explain these differences will be needed.

We found that a combination of Ser211Ala/Ser161Ala, as well as Asp25Gln/Asp256Gln or His214Phe/His286Phe, substitutions in EH_{1AB} completely arrested the activity of all 96 esters tested in this study, including the 72 esters converted by EH_{1AB}, as well as *p*NPP (see Supporting Results and Figures S9 and S10), unambiguously confirming their role in catalysis.

Altogether, these data unambiguously demonstrated that two different active sites can functionally and efficiently co-exist in a single serine ester hydrolase. However, we did not identify a single specific advantage of this engineered ester hydrolase with two active sites, albeit we found that properties such as substrate binding and the optimum pH were altered. We hypothesized why the engineered enzyme had lower performance. One of the mutations needed to introduce the new active site most likely slightly perturbs the native catalytic triad arrangement, which has negative consequences for the activity phenotype of the engineered mutant. Locating an extra binding pocket in positions not disturbing the original pocket and converting it into a catalytic site could solve this issue. In terms of substrate spectra, we should again highlight that EH_{1A} already contains an active site capable of hydrolyzing a prominent number of esters, which is significantly higher than for the other 146 natural ester hydrolases, including the best commercial prototypes such as CalB (Figure 6).³⁰ Therefore, the probability of designing an active site capable of converting an additional set of esters compared to EH_{1A} may be low. To unambiguously prove the advantage of introducing two active sites for expanding the substrate range, one should start with an ester hydrolase with a smaller substrate range and insert an extract active with different specificity or selectivity, which is actually being undertaken.

CONCLUSION

Is it possible to create enzymes with more than one active site? This idea, which has never been successfully attempted, was the starting point of the present study.

Our work describes, for first time to the best of our knowledge, a neat proof-of-concept demonstration of the use of computational approaches to discover in an ester hydrolase already containing a native catalytic site, a novel binding site where a second catalytic site can be successfully introduced to generate a protein with two functional active sites for ester hydrolysis. Both active sites were structurally and catalytically different. We did not only prove that introducing extra functional groups into a catalytic serine ester hydrolase is plausible, but also examined the catalytic potential of the resulting variant as compared to the target ester hydrolase and other reported native ester hydrolases. We confirmed the catalytic potential of the newly introduced active site as an individual catalytic entity or when combined with the native active site in the same polypeptide. We would like to highlight that a number of control experiments and analytics were performed to support our assertions and to ensure the reported catalytic activity for the de novo designed active site was not due to a contaminant (for details see Supporting Results Section).

Can every ester hydrolase be evolved for multiple reactive groups? Possibly not, as introducing multiple catalytic environments would depend on the possibility of finding an extra binding site in an enzyme of a size large enough to bind into defined cavities. Also, if found, it may not be turned into a catalytic active site after introducing the appropriate mutations because catalysis requires amino acids at appropriate distances and angles.¹¹ Extending the analysis to other ester hydrolases and enzyme scaffolds will help to clarify whether the possibility to introduce extra catalytic sites is fortuitous or could be reproducible and to what extent can it happen.

The scientific and economic benefits of introducing multiple active sites into an enzyme are enormous and may revolutionize the fields of protein design and enzymology. Thus, converting binding pockets into extra catalytic active sites with different active site cavity volumes, accessible surface areas and distances and angles between amino acids participating in the reactions compared to the native binding pockets will open novel catalytic opportunities for a given enzyme. These opportunities will come from the presence of different catalytic environments in the same polypeptide. Creating enzymes with multiple reactive groups may improve the substrate conversion range of enzymes otherwise containing only one active site, increase the competitiveness of enzymes over heterogeneous catalytic materials containing a high number of functional groups, and open application ranges with consequent reductions in production costs of other multiple enzymes. Moreover, adding new active sites to an enzyme whose original/native one cannot well be implemented for a particular application, but its overall structure has the appropriate technical requirements (scalability and robustness), may open new opportunities for a given enzyme. Such opportunities can be further explored through rational design and directed evolution focusing on residues around the new active site. Further studies in that direction would lay the foundation for

streamlined and reproducible approaches for designing enzymes with multiple active sites. It would also promote new advances in a mechanistic understanding of enzymatic principles and the evolution of catalytic environments, particularly for revealing why certain areas in protein macromolecules are favored over others for incorporating active sites.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge via the Internet at <http://pubs.acs.org>.

Supporting Results Section, Supporting Figures S1 to S11 and Supporting Tables S1 to S4 (PDF).

AUTHOR INFORMATION

Corresponding Author

*victor.guallar@bsc.es

*mferrer@icp.csic.es

ORCID

Gerard Santiago: 0000-0002-0506-3049

Peter N. Golyshin: 0000-0002-5433-0350

Víctor Guallar: 0000-0002-4580-1114

Manuel Ferrer: 0000-0003-4962-4714

Author Contributions

^oThese authors contributed equally to this work.

Funding Sources

This project received funding from the European Union's Horizon 2020 research and innovation program [Blue Growth: Unlocking the potential of Seas and Oceans] under grant agreement no. [634486] (project acronym INMARE). This research was also supported by the grants PCIN-2014-107 (within ERA NET IB2 grant nr. ERA-IB-14-030 - MetaCat), PCIN-2017-078 (within the ERA-MarineBiotech grant ProBone), BIO2014-54494-R, CTQ2016-79138-R and BIO2017-85522-R from the Spanish Ministry of Economy, Industry and Competitiveness. P.N.G. gratefully acknowledges funding from the UK Biotechnology and Biological Sciences Research Council (grant no. BB/M029085/1). R.B. and P.N.G. acknowledge the support of the Supercomputing Wales project, which is part-funded by the European Regional Development Fund (ERDF) via Welsh Government. P.N.G. acknowledges the support of the Centre of Environmental Biotechnology Project funded by the European Regional Development Fund (ERDF) through Welsh Government. The authors gratefully acknowledge financial support provided by the European Regional Development Fund (ERDF). The MALDI-TOF/TOF analysis was performed in the proteomics facility of the Spanish National Center for Biotechnology (CNB-CSIC) that belongs to ProteoRed, PRB2-ISCIII, supported by grant PT13/0001.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

C.C. thanks the Spanish Ministry of Economy, Industry and Competitiveness for a PhD fellowship (Grant BES-2015-

073829. Author would like to acknowledge Sergio Ciordia and María C. Mena for MALDI-TOF/TOF analysis.

REFERENCES

- (1) Dufaud, V., and Davis, M. E. (2003) Design of heterogeneous catalysts via multiple active site positioning in organo-inorganic hybrid materials. *J. Am. Chem. Soc.* 125, 9403-9413.
- (2) Jing, W., and DeAngelis, P.L. (2000) Dissection of the two transferase activities of the *Pasteurella multocida* hyaluronan synthase: two active sites exist in one polypeptide. *Glycobiology* 10, 883-889.
- (3) Frickel, E. M., Jemth, P., Widersten, M., and Mannervik, B. (2001) Yeast glyoxalase I is a monomeric enzyme with two active sites. *J. Biol. Chem.* 276, 1845-1849.
- (4) Deponte, M., Sturm, N., Mittler, S., Harner, M., Mack, H., and Becker, K. (2007) Allosteric coupling of two different functional active sites in monomeric *Plasmodium falciparum* glyoxalase I. *J. Biol. Chem.* 282, 28419-28430.
- (5) Khare, S. D., Kipnis, Y., Greisen, P. Jr., Takeuchi, R., Ashani, Y., Goldsmith, M., Song, Y., Gallaher, J. L., Silman, I., Leader, H., Sussman, J.L., Stoddard, B.L., Tawfik, D. S., and Baker, D. (2012) Computational redesign of a mononuclear zinc metalloenzyme for organophosphate hydrolysis. *Nat. Chem. Biol.* 8, 294-300.
- (6) Seelig, B., and Szostak, J. W. (2007) Selection and evolution of enzymes from a partially randomized non-catalytic scaffold. *Nature* 448, 828-831.
- (7) Toscano, M. D., Woycechowsky, K. J., Hilvert, D. (2007) Minimalist active-site redesign: teaching old enzymes new tricks. *Angew. Chem. Int. Ed. Engl.* 46, 3212-3236.
- (8) Jiang, L., Althoff, E. A., Clemente, F. R., Doyle, L., Röthlisberger, D., Zanghellini, A., Gallaher, J. L., Betker, J. L., Tanaka, F., Barbas, C. F. 3rd., Hilvert, D., Houk, K. N., Stoddard, B. L., and Baker, D. (2008) De novo computational design of retro-aldol enzymes. *Science* 319, 1387-1391.
- (9) Korendovych, I. V., Kulp, D. W., Wu, Y., Cheng, H., Roder, H., and DeGrado, W. F. (2011) Design of a switchable eliminase. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6823-6827.
- (10) Der, B. S., Edwards, D. R., and Kuhlman, B. (2012) Catalysis by a de novo zinc-mediated protein interface: implications for natural enzyme evolution and rational enzyme engineering. *Biochemistry* 51, 3933-3940.
- (11) Richter, F., Blomberg, R., Khare S. D., Kiss, G., Kuzin, A. P., Smith, A. J., Gallaher J., Pianowski, Z., Helgeson, R. C., Grjasnow, A., Xiao, R., Seetharaman, J., Su, M., Vorobiev, S., Lew, S., Forouhar, F., Kornhaber, G. J., Hunt, J. F., Montelione, G. T., Tong, L., Houk, K. N., Hilvert, D., and Baker, D. (2012) Computational design of catalytic dyads and oxyanion holes for ester hydrolysis. *J. Am. Chem. Soc.* 134, 16197-16206.
- (12) Zastrow, M. L., Peacock, A. F., Stuckey, J. A., and Pecoraro, V. L. (2012) Hydrolytic catalysis and structural stabilization in a designed metalloprotein. *Nat. Chem.* 4, 118-123.
- (13) Korendovych, I. V., and DeGrado, W. F. (2014) Catalytic efficiency of designed catalytic proteins. *Curr. Opin. Struct. Biol.* 27, 113-121.
- (14) Rufo, C. M., Moroz, Y. S., Moroz, O. V., Stöhr, J., Smith, T. A., Hu, X., DeGrado, W. F., and Korendovych, I. V. (2014) Short peptides self-assemble to produce catalytic amyloids. *Nat. Chem.* 6, 303-309.
- (15) Raymond, E. A., Mack, K. L., Yoon, J. H., Moroz, O. V., Moroz, Y. S., and Korendovych, I. V. (2015) Design of an allosterically regulated retroaldolase. *Protein Sci.* 24, 561-570.
- (16) Moroz, Y. S., Dunston, T. T., Makhlynets, O. V., Moroz, O. V., Wu, Y., Yoon, J. H., Olsen, A. B., McLaughlin, J. M., Mack, K. L., Gosavi, P. M., van Nuland, N. A., and Korendovych, I. V. (2015) New tricks for old proteins: single mutations in a nonenzymatic protein give rise to various enzymatic activities. *J. Am. Chem. Soc.* 137, 14905-14911.
- (17) Aranda, J., Cerqueira, N. M., Fernandes, P. A., Roca, M., Tuñón, I., and Ramos, M.J. (2014) The catalytic mechanism of carboxylesterases: a computational study. *Biochemistry* 53, 5820-5829.
- (18) Ferrer, M., Bargiela, R., Martínez-Martínez, M., Mir, J., Koch, R., Golyshina, O. V., and Golyshin, P. N. (2015) Biodiversity for biocatalysis: A review of the α/β -hydrolase fold superfamily of esterases-lipases discovered in metagenomes. *Biocatal. Biotransform.* 33, 235-249.
- (19) Zastrow, M. L., and Pecoraro, V. L. (2013) Influence of active site location on catalytic activity in de novo-designed zinc metalloenzymes. *J. Am. Chem. Soc.* 135, 5895-5903.
- (20) Ross, M. R., White, A. M., Yu, F., King, J. T., Pecoraro, V. L., and Kubarych, K. J. (2015) Histidine orientation modulates the structure and dynamics of a de novo metalloenzyme active site. *J. Am. Chem. Soc.* 137, 10164-10176.
- (21) Borrelli, K. W., Cossins, B., and Guallar, V. (2010) Exploring hierarchical refinement techniques for induced fit docking with protein and ligand flexibility. *J. Comput. Chem.* 31, 1224-1235.
- (22) Hernández-Ortega, A., Borrelli, K., Ferreira, P., Medina, M., Martínez, A. T., Guallar, V. (2011) Substrate diffusion and oxidation in GMC oxidoreductases: an experimental and computational study on fungal aryl-alcohol oxidase. *Biochemical J.* 436, 341-350.
- (23) Carlson, H. A., Smith, R. D., Damm-Ganamet, K. L., Stuckey, J. A., Ahmed, A., Convery, M. A., Somers, D. O., Kranz, M., Elkins, P. A., Cui, G., Peishoff, C. E., Lambert, M. H., and Dunbar, J. B. Jr. (2016) CSAR 2014: a benchmark exercise using unpublished data from pharma. *J. Chem. Inf. Model.* 56, 1063-1077.
- (24) Santiago, G., de Salas, F., Lucas, M.F., Monza, E., Acebes, S., Martínez, A. T., Camarero, S., and Guallar, V. (2016) Computer-aided laccase engineering: toward biological oxidation of arylamines. *ACS Catalysis* 6, 5415-5423.
- (25) Lecina, D., Gilabert, J. F., and Guallar, V. (2017) Adaptive simulations, towards interactive protein-ligand modeling. *Sci. Rep.* 7, 8466.
- (26) Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., and Jorgensen, W. L. (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* 105, 6474-6487.
- (27) Bochevarov, A. D. (2013) Quantum mechanics approaches to drug research in the era of structural chemogenomics. *Int. J. Quantum Chem.* 113, 2110-2142.
- (28) Shan, Y., Kim, E. T., Eastwood, M. P., Dror, R. O., Seeliger, M. A., and Shaw, D. E. (2011) How does a drug molecule find its target binding site? *J. Am. Chem. Soc.* 133, 9181-9183.
- (29) Martínez-Martínez, M., Alcaide, M., Tchigvintsev, A., Reva, O., Polaina, J., Bargiela, R., Guazzaroni, M.E., Chicote, A., Canet, A., Valero, F., Rico Eguizabal, E., Guerrero, Mdel C., Yakunin, A. F., and Ferrer, M. (2013) Biochemical diversity of carboxyl esterases and lipases from Lake Arreo (Spain): a metagenomic approach. *Appl. Environ. Microbiol.* 79, 3553-3562.
- (30) Martínez-Martínez, M., Coscolín, C., Santiago, G., Chow, J., Stogios, P., Bargiela, R., Gertler, C., Navarro-Fernández, J., Bollinger, A., Thies, S., Méndez-García, C., Popovic, A., Brown, G., Chernikova, T. N., García-Moyano, A., Bjerga, G. E. K., Pérez-García, P., Hai, T., Del Pozo, M. V., Stokke,

- R., Steen, I. H., Cui, H., Xu, X., Nocek, B., Alcaide, M., Distaso, M., Mesa, V., Peláez, A. I., Sánchez, J., Buchholz, P. C. F., Pleiss, J., Fernández-Guerra, A. F., Glöckner, F. O., Golyshina, O. V., Yakimov, M. M., Savchenko, A., Jaeger, K-E., Yakunin, A. F., Streit, W. R., Golyshin, P. N., Guallar, V., and Ferrer, M. (2018) Determinants and prediction of esterase substrate promiscuity patterns. *ACS Chem. Biol.* 13, 225-234.
- (31) Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.
- (32) Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248-254.
- (33) Janes, L. E., Löwendahl, C., and Kazlauskas, R. J. (1998) Rapid quantitative screening of hydrolases using pH indicators. Finding enantioselective hydrolases. *Chem. Eur. J.* 4, 2317-2324.
- (34) Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R., and Sherman, W. (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* 27, 221-234.
- (35) Shivakumar, D., Williams, J., Wu, Y., Damm, W., Shelley, J., and Sherman, W. J. (2010) Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. *Chem. Theory Comput.* 6, 1509-1519.
- (36) Schevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal. Chem.* 68, 850-858.
- (37) Berg, J.M., Tymoczko, J.L., and Stryer, L. (2002) *Biochemistry*. 5th edition. New York: W H Freeman; Section 8.4. The Michaelis-Menten model accounts for the kinetic properties of many enzymes. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK22430/>

Table of Contents artwork

EH_{1AB} structure with the two active sites highlighted, in red the original active site and in blue the engineered one. EH_{1AB} structure has been created with Maestro.

